JOURNAL OF BACTERIOLOGY, June 1992, p. 3807–3811
0021-9193/92/113807-05$02.00/0

Vol. 174, No. 11

NOTES

Dissection of the *Salmonella typhimurium* Genome by Use of a Tn*5* Derivative Carrying Rare Restriction Sites

KWONG KWOK WONG* AND MICHAEL McCLELLAND

California Institute of Biological Research, 11099 North Torrey Pines Road, La Jolla, California 92037

Received 16 September 1991/Accepted 21 March 1992

A polylinker with rare restriction sites was introduced into a mini-Tn*5* derivative. These sites include M.*Xba*I-*Dpn*I (TCTAGATCTAGA), which is rare in most bacterial genomes, *Swa*I (ATTTAAAT) and *Pac*I (TTAATTAA), which are rare in G+C-rich genomes, *Not*I (GCGGCCGC) and *Sfi*I (GGCCN$_5$GGCC), which are rare in A+T-rich genomes, and *Bln*I (CCTAGG), *Spe*I (ACTAGT), and *Xba*I (TCTAGA), which are rare in the genomes of many gram-negative bacteria. This Tn*5*(pfm) (pulsed-field mapping) transposon carries resistance to chloramphenicol and kanamycin to allow selection in a wide variety of background genomes. This Tn*5*(pfm) was integrated randomly into the *Salmonella typhimurium* and *Serratia marcescens* genomes. Integration of the new rare *Swa*I, *Pac*I, *Bln*I, *Spe*I, and *Xba*I sites was assayed by restriction digestion and pulsed-field gel electrophoresis. Tn*5*(pfm) constructs could be valuable tools for pulsed-field mapping of gram-negative bacterial genomes by assisting in the production of physical maps and restriction fragment catalogs. For the first applications of a Tn*5*(pfm), we bisected five of the six largest *Bln*I fragments in the *S. typhimurium* genome, bisected the linearized 90-kb pSLT plasmid, and used Tn*5*(pfm) and Tn*10* to trisect the largest *Bln*I fragment.

Transposons carrying rare DNA cleavage sites could be used in the physical and genetic mapping of bacterial genomes by pulsed-field gel electrophoresis (PFGE) (12, 25). Some transposons naturally carry restriction sites that are rare in some genomes. For example, Tn*5*, which can transpose in most gram-negative bacteria (1, 2), has a *Not*I (GCGGCCGC) site which occurs only 22 times in the *Escherichia coli* K-12 genome (24, 25). *Not*I sites are rare in most A+T-rich genomes (17). Tn*10*, which transposes in some enterobacteria, naturally carries a *Bln*I (or *Avr*II) (CCTA GG) site, which is rare in many enterobacterial genomes (3, 4, 11, 12, 29). Tn*917*, which transposes in some gram-positive bacteria, has been engineered to carry a *Sma*I site (32), which is rare in very A+T-rich genomes such as the *Staphylococcus aureus* genome (20, 21, 28). However, no transposon carries a battery of restriction sites, some of which are likely to be rare in any bacterial genome.

We constructed a series of such transposons based on Tn*5* because this transposon can be used in most gram-negative species (1, 2). We used the delivery system developed by the Timmis laboratory (5) in which the Tn*5* transposon is on a mobilizable plasmid with an R6K origin of replication that cannot function in the absence of its replication initiator protein. The transposase gene is adjacent to but outside the transposon. The plasmid contains the conjugal transfer origin (*oriT* sequence) of the RP4 plasmid. The *tra* functions and the R6K π protein (required for initiation of plasmid replication) are provided by a λ*pir* lysogen of the *E. coli* host S17-1 (15, 23). Because the plasmid is unable to replicate in the absence of *pir*, the transposase is quickly lost when the plasmid is transferred to the new bacterium, so that any transposition that may have occurred to the recipient genome is stabilized.

Construction of the Tn*5* derivatives was performed as follows. Two complementary 44-base oligonucleotides were manufactured by Genosys (Houston, Tex.), 5′ TCGATT TAAA TCTAGATCTA GATCTAGATT AATTAAGCGG CCGC 3′ and 5′ TCGAGCGGCC GCTTAATTAA TCTA GATCTA GATCTAGATT TAAA 3′. The duplex formed by annealing these two oligodeoxyribonucleotides carries sites for *Swa*I, M.*Xba*I-*Dpn*I (a combination of the *Xba*I methylase and the methylation-dependent restriction endonuclease *Dpn*I), *Pac*I, and *Not*I. The duplex was inserted into the

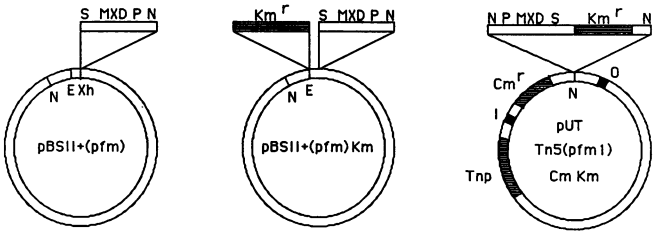

FIG. 1. Construction of a Tn*5*(pfm1) Cm Km derivative carrying multiple rare cleavage sites. A synthetic 44-bp duplex that carries sites for *Swa*I (S), M.*Xba*I-*Dpn*I (MXD), *Pac*I (P), and *Not*I (N) was inserted into the *Xho*I (Xh) site of pBSKII+ (Stratagene). An *Eco*RI (E) fragment carrying a Km^r gene of Tn*903* was ligated into the *Eco*RI site of the purified plasmid pBSII(pfm1+). The resulting plasmid, pBSII(pfm1+) Km, carrying Km^r and Ap^r genes was purified and cleaved with *Not*I (N). This released a large fragment carrying the Km^r gene and most of the pBSKII+ polylinker. This fragment was cloned into the unique *Not*I site of a pUT derivative carrying mini-Tn*5* Cm and the transposase (Tnp). I and O are the two minimal IS*50* ends required for transposition. Tn*5*(pfm1) Cm Km also carried two *Sfi*I and two *Bln*I sites.

* Corresponding author.

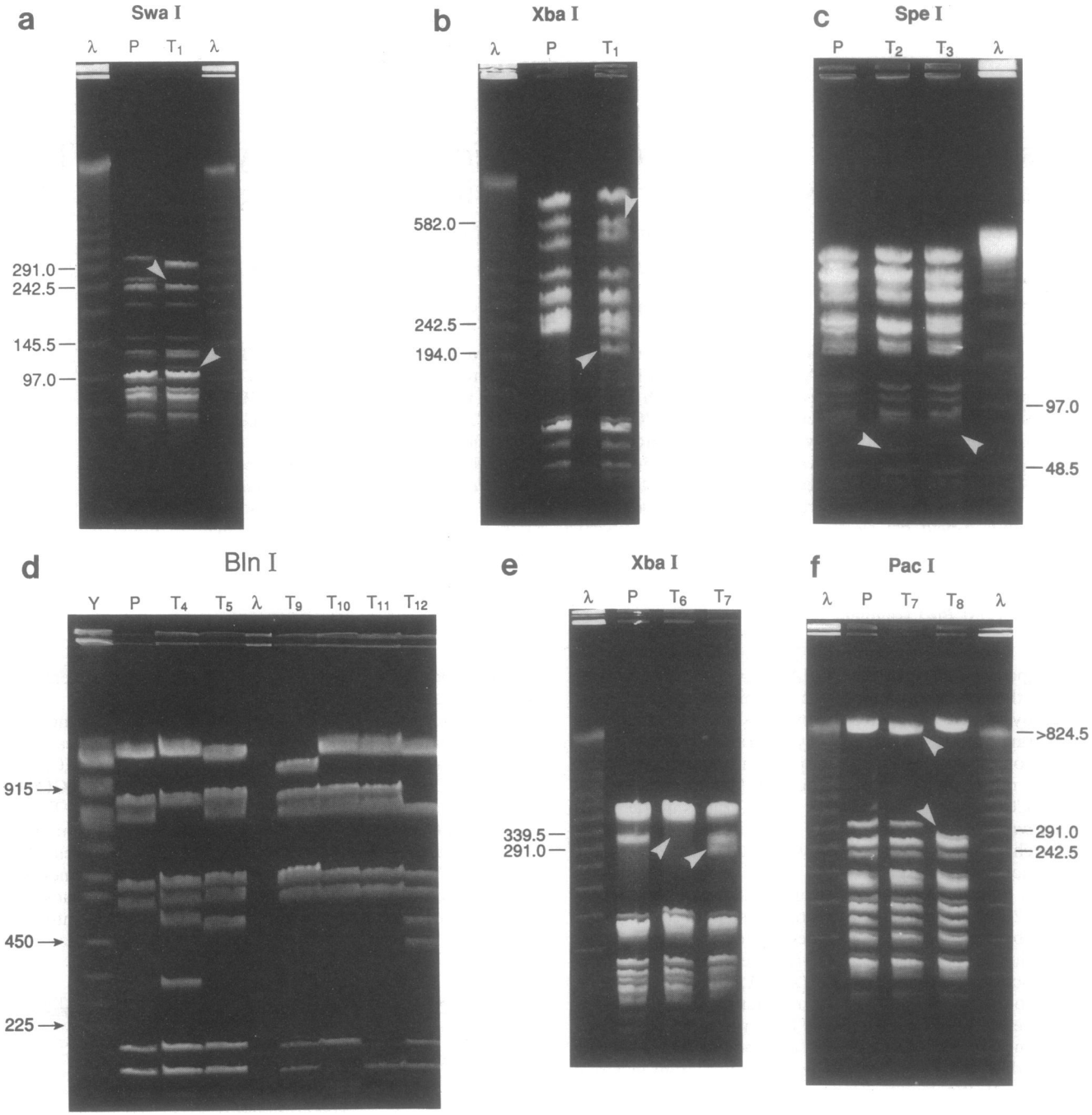

FIG. 2. New rare cleavage sites observed by PFGE. Agarose plugs of genomic DNA from the *S. typhimurium* LT2 and *S. marcescens* parental and Tn*5*(pfm) Cm Km insertions were cleaved with 10 U of *Bln*I, *Spe*I, *Swa*I, *Xba*I, or *Pac*I. The DNA fragments were resolved on a transverse alternating-field electrophoresis (Beckman Instruments) PFGE apparatus. Lane λ, bacteriophage λ concatemers (48,502-bp unit length); lane Y, yeast chromosome markers; lane P, parental genomic DNA without any inserted transposon. The other lanes, indicated by T_n in each panel, are a selection of genomic DNAs with Tn*5*(pfm) Cm Km insertions. Arrowheads indicate some of the changes in restriction pattern. Lanes T_1 through T_5 and T_9 through T_{12} contain *S. typhimurium* strains and lanes T_6 through T_8 contain *S. marcescens* strains. In panel d, the strains shown in lanes T_4, T_5, T_9, T_{10}, T_{11}, and T_{12} have Tn*5*(pfm) insertions in *Bln*I fragments previously designated C, A, A, G, F1, and B, respectively (29). DNA sizes are given in kilobases.

*Xho*I site of pBSKII (Stratagene, La Jolla, Calif.), and the resulting clones in each orientation, pBS(pfm1+) and pBS (pfm1−), were confirmed by DNA sequencing. An *Eco*RI fragment carrying the kanamycin resistance (Km^r) gene of Tn*903* (19) (Pharmacia, Piscataway, N.J.) was ligated into the *Eco*RI site of the purified plasmid pBS(pfm1+) carrying the rare cleavage site cassette. A resulting plasmid, pBS (pfm1+) Km, carrying Km^r and ampicillin resistance (Ap^r) genes, was purified and cleaved with *Not*I. This released a large fragment carrying the Km^r gene and the rare cleavage

TABLE 1. Location of 17 mapped Tn*5*(pfm) insertions

Strain	*Bln*I fragment[a]	Approximate location (min)[b]
KKW1001	D	65 or 67
KKW1002	C	31
KKW1003	A	8
KKW1004	C	30
KKW1005	A	10
KKW1006	A	95 or 19
KKW1007	C	32
KKW1008	C	22
KKW1009	A	6
KKW1010	C	27
KKW1011	A	1 or 15
KKW1012	A	14
KKW1013	None (pSLT)	None (pSLT)
KKW1014	F1	59 or 59
KKW1015	B	44
KKW1016	C	31
KKW1017	C	32

[a] Letters indicate *Bln*I restriction fragments in the published map (29).
[b] Each Tn*5*(pfm) insertion is first mapped by PFGE to two possible locations relative to the ends of the naturally occurring *Bln*I fragment bisected by the insertion. Transduction of the Tn*5*(pfm) insertion into a strain carrying a genetically mapped Tn*10* insertion in the same *Bln*I fragment allows Tn*5*(pfm) insertions to be mapped to a single location by measuring the distance to the Tn*10* insertion by PFGE. For instance, those insertions in fragment C were mapped with *pyrF*::Tn*10*.

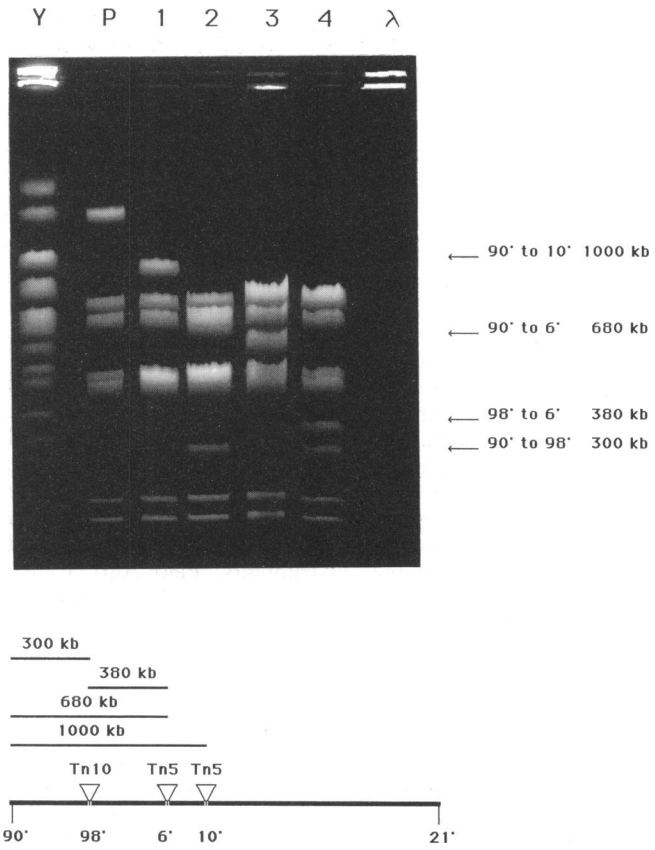

FIG. 3. Drop out of the region between two transposon insertions. Tn*5*(pfm) insertions physically mapped to 6 and 10 min in the *S. typhimurium* genome were each transduced into a strain carrying Tn*10* in *pyrB* at 98 min. In each case, genomic DNA was prepared and digested with *Bln*I as in Fig. 2. The region between the transposon-introduced *Bln*I sites is a new 560-kb fragment that comigrates with another fragment in one case and a well-resolved 380-kb fragment in the other case. Lane Y, yeast chromosome markers; lane P, parental genomic DNA without any inserted transposon; lane λ, bacteriophage λ concatemers (48,502-bp unit length); lane 1, Tn*5*(pfm) inserted at 10 min; lane 2, Tn*5*(pfm) inserted at 10 min and *pyrB*::Tn*10* inserted at 98 min; lane 3, Tn*5*(pfm) inserted at 6 min; lane 4, Tn*5*(pfm) inserted at 6 min and *pyrB*::Tn*10* inserted at 98 min.

site cassette with *Swa*I, *Pac*I, *Not*I, M.*Xba*I-*Dpn*I, and most of the pBSKII+ polylinker, including sites for *Spe*I, *Eag*I, *Bam*HI, *Sma*I, *Pst*I, *Eco*RI, *Eco*RV, *Hin*dIII, *Cla*I, *Sal*I, and *Dra*I. This fragment was cloned into the unique *Not*I site of a pUT derivative carrying a chloramphenicol-resistant (Cm^r) mini-Tn*5* (5). Cm^r Km^r transformants were selected, purified, and analyzed by restriction digestion. Of the two possible orientations of insertion, only one was obtained, referred to as Tn*5*(pfm1) Cm Km, where pfm is an abbreviation for pulsed-field mapping and 1 indicates the orientation of the Km^r cassette. The construct carried the expected restriction sites (data not shown). Tn*5*(pfm1) Cm Km also carried two *Sfi*I and two *Bln*I sites. *Bln*I was purchased from Takara Biochemical Inc. (Berkeley, Calif.), *Pac*I was from New England Biolabs (Beverly, Mass.), and *Swa*I was from Boehringer Mannheim Biochemicals (Indianapolis, Ind.). All other enzymes were purchased from Stratagene. The plasmids are diagrammed in Fig. 1.

An indication of which restriction cleavage sites are likely to be rare in a particular genome can be gathered simply on the basis of the G+C content of the genome. If a few thousand base pairs of genomic sequence are available, more accurate calculations can be obtained by using di- and trinucleotide frequencies (17). Theoretical calculations using base composition alone indicate that *Not*I and *Sfi*I should typically occur less than once every 10^6 bp in a genome that contains greater than 65% A+T. For example, *Not*I and *Sfi*I sites do not occur at all in the *S. aureus* genome which contains 2.8×10^6 bp and is 68% A+T (28). Similarly, *Pac*I and *Swa*I sites should occur less than once every 10^6 bp in most genomes that have a G+C content over 65%. For example, the 6-megabase *Rhizobium meliloti* genome (67% G+C) has three large replicons that contain only four *Pac*I sites and six *Swa*I sites (26). The M.*Xba*I-*Dpn*I site is sufficiently long that it should be very rare or absent in most bacterial genomes.

*Bln*I, *Spe*I, and *Xba*I recognize target sites that are only 6 bp long. Nevertheless, these sites are extremely rare in some gram-negative bacteria, as diverse as *E. coli* and *R. meliloti*, because the tetranucleotide CTAG is underrepresented about 10-fold in these genomes relative to predictions made on the basis of base composition (3, 11, 12). For instance, *Bln*I (or *Avr*II) sites occur less than 15 times in most *E. coli* K-12 genomes (4). Similarly, *Bln*I sites are the rarest of any known restriction endonuclease in the *Salmonella typhimurium* LT2 genome; they occur only about 12 times, of which at least five sites may occur in or near 16S rRNA genes. The physical map of *Bln*I sites in the *S. typhimurium* genome has been completed (29).

We checked the conjugation and transposition of Tn*5*(pfm) in *S. typhimurium* LT2 and *Serratia marcescens*, because they are known to be good recipients for conjugation and Tn*5* transposition (1). Also, *Bln*I, *Spe*I, and *Xba*I sites should be relatively rare in these enterobacteria, and *Pac*I and *Swa*I should be relatively rare in the genome of *S. marcescens* (59% G+C) (18). The Tn*5*(pfm) Cm Km trans-

poson was conjugated (14) into a tetracycline-resistant (Tc^r) *S. typhimurium* LT2 and a rifampin-resistant (Rf^r) *S. marcescens*. Exconjugants were selected by using 20 μg of chloramphenicol per ml with either 20 μg of tetracycline per ml or 100 μg of rifampin per ml used to counterselect against the donor. The transfer allowed a substantial number of transposition events to be obtained, but the frequency was about 10-fold lower than that found in experiments with wild-type Tn*5*. Ninety-five percent of Cm^r recipients were Ap^s, indicating they had lost the plasmid and transposition to a stable replicon had occurred. Unexpectedly, 5% of *S. typhimurium* exconjugants remain unstably resistant to ampicillin. The plasmid perhaps carries a cryptic origin that allows very low-level replication in *S. typhimurium*.

Ap^s Cm^r *S. typhimurium* and *S. marcescens* exconjugants were used to prepare genomic DNA in agarose (InCert; FMC, Rockland, Maine) plugs (9). The DNAs were cleaved with 10 U of *Bln*I, *Xba*I, *Spe*I, or *Swa*I in 1× universal buffer (as defined in reference 8 except that potassium acetate replaces potassium glutamate) for 6 h at 37°C. The DNA fragments were resolved on a transverse alternating-field electrophoresis (Beckman Instruments, Palo Alto, Calif.) PFGE apparatus (7, 22) using 0.5× TAE and 1% LE agarose (FMC). Pulse times were 5, 10, 15, 25, and 50 s for 6 h each, except for the *Bln*I digest which was separated using pulse times of 10, 30, 60, 120, and 240 s for 0.5, 2, 15, 10, and 8 h, respectively. In all cases, differences in the restriction pattern consistent with integration of a single Tn*5*(pfm) were noted with at least one enzyme digest, as can be seen in Fig. 2, for example. These observations indicate that the transposons are functional and carry the rare cleavage site cassette after transposition. The introduction of *Not*I and *Sfi*I sites (also present in the original constructs [5]) was not tested as these sites are very common in the genomes of *S. typhimurium* (50% G+C) and *S. marcescens* (55% G+C). M.*Xba*I-*Dpn*I should also be cleaved after methylation in vivo (9 and data not shown). As a first application, Fig. 2d presents Tn*5* insertions into five of the nine *Bln*I fragments from the *Salmonella* genome resolved by PFGE. The insertions shown include one in the 90-kb pSLT megaplasmid (13), shown in Fig. 2d, lane T_{10}. These insertions allow us to further subdivide the *Salmonella* genome on the basis of the completed *Bln*I restriction map (29). Eleven other characterized insertions are not shown. Such fragments are a source of genomic subchromosomal clone libraries and dot blot arrays for physical mapping of genes (30).

Other DNA cleavage sites that are likely to be absent or rare in bacterial genomes might be engineered into transposons. Examples include the *Lac*I-methylase-endonuclease "Achilles heel" combination (10), the 18-bp *Sce*I site from *Saccharomyces cerevisiae* (16), or triplex helix cleavage sites (27). In addition, Tn*5*(pfm) transposons could be developed that carry plasmid origins, allowing the region around the insertion site to be cloned, as has been described for Tn*5* (2) and the omegon transposon (6). Tn*5*(pfm1) derivatives that carry other resistance markers can also be developed.

One potentially important application of Tn*5*(pfm) is the introduction of two very rare or otherwise absent restriction sites into a bacterial genome using separate insertions with different selectable markers. The region between the transposons can be excised using restriction digestion and PFGE, a method we term "drop out". In one of its most useful manifestations, two transposons with different selectable markers are used to make strains, each carrying a transposon at a different place in the genome. Transposons integrated into different locations in the genome can then be placed in the same strain background by conjugation or transduction, followed by selection for both markers. The region between the two different transposons can then be dropped out. Because the transposon integrations can be combined in a matrix of pairwise combinations, one can systematically develop a set of strains, each carrying two transposons, that divide the genome into a set of overlapping fragments. For the first application, we placed Tn*5*(pfm) at a variety of positions in the *S. typhimurium* genome and transduced them into a background genome containing Tn*10* insertions that also carry the rare *Bln*I site (29). A list of some mapped Tn*5*(pfm) insertions is given in Table 1. For example, a Tn*5*(pfm) at 6 min in the largest naturally occurring *Bln*I fragment was placed in a background with a Tn*10* at 98 min. This enabled this fragment to be trisected and drop out a 380-kb fragment (Fig. 3, lane 4). Among the interesting loci on this new fragment is the putative location of the *trpR* gene which has proved difficult to clone from *S. typhimurium* (31). This fragment and a variety of others produced in a similar manner have been purified as part of the systematic production of subchromosomal libraries and a complete dot blot array (30).

We thank Ken Timmis for the mini-Tn*5* constructs and strains, John Roth and Ken Sanderson for *S. typhimurium* strains, Alan Greener for the *S. marcescens* strains and *tra*-overproducing *E. coli* strains, and Rhonda Honeycutt, Stan Maloy, Alan Greener, Bruno Sobral, and John Welsh for many helpful discussions and critical reading of the manuscript. We thank C. Arsmith for inspiration.

This work is funded in part by a grant to M.M. from NIH and to Bruno Sobral from the U.S. Department of Agriculture.

REFERENCES

1. **Berg, D. E.** 1989. Transposon Tn*5*, p. 185–210. *In* D. E. Berg and M. Howe (ed.), Mobile DNA. American Society for Microbiology, Washington, D.C.
2. **Berg, C. M., D. E. Berg, and E. A. Groisman.** 1989. Transposable elements and genetic engineering of bacteria, p. 879–926. *In* D. E. Berg and M. Howe (ed.), Mobile DNA. American Society for Microbiology, Washington, D.C.
3. **Bhagwat, A. S., and M. McClelland.** DNA mismatch correction by very short patch repair may have altered the abundance of oligonucleotides in the *E. coli* genome. Nucleic Acids Res., in press.
4. **Daniels, D. L.** 1990. The complete *Avr*II restriction map of the *E. coli* genome and comparisons of several laboratory strains. Nucleic Acids Res. **18:**2649–2651.
5. **de Lorenzo, V., M. Herrero, U. Jakubzik, and K. N. Timmis.** 1990. Mini-Tn*5* transposon derivatives for insertion mutagenesis, promoter probing, and chromosomal insertion of cloned DNA in gram-negative eubacteria. J. Bacteriol. **172:**6568–6572.
6. **Fellay, R., H. M. Krisch, P. Prentki, and J. Frey.** 1989. Omegon-Km, a transposable element designed for *in vivo* insertional mutagenesis and cloning of genes in gram-negative bacteria. Gene **76:**215–226.
7. **Gardiner, K., W. Laas, and D. Patterson.** 1986. Fractionation of large mammalian DNA restriction fragments using vertical pulsed-field gradient gel electrophoresis. Somatic Cell Mol. Genet. **12:**185–195.
8. **Hanish, J., and M. McClelland.** 1988. Activity of restriction endonucleases and methylases in potassium glutamate buffers (KGB). Gene Anal. Tech. **5:**105–107.
9. **Hanish, J., and M. McClelland.** 1991. Enzymatic cleavage of a bacterial chromosome at an inserted 12-base-pair recognition site. Nucleic Acids Res. **19:**829–832.
10. **Koob, M., E. Grimes, and W. Szybalski.** 1988. Conferring operator specificity on restriction endonucleases. Science **241:** 1084–1086.
11. **McClelland, M., and A. S. Bhagwat.** 1992. Biased DNA repair. Nature (London) **355:**596–597.
12. **McClelland, M., R. Jones, Y. Patel, and M. Nelson.** 1987.

Restriction endonucleases for pulsed field mapping of bacterial genomes. Nucleic Acids Res. **15**:5985–6005.

13. **Michiels, T., M. Y. Popoff, S. Duriaux, C. Coynault, and G. Cornelis.** 1987. A new method for the physical and genetic mapping of large plasmids: application to the localization of the virulence determinants of the 90kd plasmid of *Salmonella typhimurium*. Microb. Pathog. **3**:109–116.
14. **Miller, J. H.** 1972. Experiments in molecular genetics, p. 82–84. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
15. **Miller, V. L., and J. J. Mekalanos.** 1988. A novel suicide vector and its use in construction of insertion mutations: osmoregulation of outer membrane proteins and virulence determinants in *Vibrio cholera* requires *toxR*. J. Bacteriol. **170**:2575–2583.
16. **Monteilhet, C., A. Perrin, A. Thierry, L. Colleaux, and B. Dujon.** 1990. Purification and characterization of the in vitro activity of *Sce*-I. Nucleic Acids Res. **18**:1407–1414.
17. **Nelson, M., and M. McClelland.** 1987. Enhancement of the apparent specificities of restriction endonucleases: applications to megabase mapping, p. 257–282. *In* J. G. Chirikjian (ed.), Gene amplification and analysis, vol. V. Elsevier, New York.
18. **Nomore, W. M., and J. R. Brown.** 1970. Guanine plus cytosine (G+C) composition of bacteria, p. H24–H74. *In* H. A. Sober (ed.), Handbook of biochemistry: selected data for molecular biology, 2nd ed. The Chemical Rubber Co., Cleveland.
19. **Oka, A., H. Sugisaki, and M. Takanami.** 1981. Nucleotide sequence of the kanamycin resistance transposon Tn*903*. J. Mol. Biol. **147**:217–226.
20. **Pattee, P.** 1990. *Staphylococcus aureus*, p. 1123–1127. *In* S. J. O'Brian (ed.), Genetic maps: locus maps of complex genomes, 5th ed. Cold Spring Harbor Laboratory Press, Plain View, N.Y.
21. **Poddar, S., and M. McClelland.** 1991. Restriction fragment fingerprints and genome sizes of *Staphylococcus* species determined using pulsed field gel electrophoresis. DNA Cell Biol. **10**:663–669.
22. **Schwartz, D. C., and C. R. Cantor.** 1984. Separation of yeast *Saccharomyces cerevisiae* chromosome sized DNA by pulsed field gradient gel electrophoresis. Cell **37**:67–75.
23. **Simon, R., U. Priefer, and A. Puhler.** 1983. A broad host range mobilization system for in vivo genetic engineering: transposon mutagenesis in gram-negative bacteria. Bio/Technology **1**:784–791.
24. **Smith, C. L., J. Econome, A. Schutt, S. Klco, and C. R. Cantor.** 1987. A physical map of the *E. coli* genome. Science **236**:1448–1453.
25. **Smith, C. L., and R. D. Kolodner.** 1988. Mapping of *E. coli* chromosomal Tn*5* and F insertions by pulsed field gel electrophoresis. Genetics **119**:227–236.
26. **Sobral, B. W. S., R. J. Honeycutt, A. G. Atherly, and M. McClelland.** 1991. Electrophoretic separation of the three *Rhizobium meliloti* replicons. J. Bacteriol. **173**:5173–5180.
27. **Strobel, S. A., and P. B. Dervan.** 1991. Single-site enzymatic cleavage of yeast genomic DNA mediated by triple helix formation. Nature (London) **350**:172–174.
28. **Weil, M., and M. McClelland.** 1989. Enzymatic cleavage of a bacterial genome into two pieces at a ten-base-pair recognition site. Proc. Natl. Acad. Sci. USA **86**:51–55.
29. **Wong, K. K., and M. McClelland.** 1992. A *Bln*I restriction map of the *Salmonella typhimurium* genome. J. Bacteriol. **174**:1656–1661.
30. **Wong, K. K., and M. McClelland.** Unpublished data.
31. **Youderian, P.** Personal communication.
32. **Youngman, P., J. B. Perkins, and R. Losick.** 1984. Construction of a cloning site near one end of Tn*917* into which foreign DNA may be inserted without affecting transposition in *Bacillus subtilis* or expression of the transposon-borne *erm* gene. Plasmid **12**:1–9.